

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: June 23, 2003, 15:11:47 ; Search time 40 Seconds  
(without alignments)  
442.218 Million cell updates/sec

Title: AAK91826

Perfect score: 965

Sequence: 1 MRKGRSLRGHDAPAPPCV.....ATELGSTELVTKTAGPEQQ 184

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

a1 number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	124	12.8	416	1 SKXLAG	dermal gland prote
2	119	12.3	552	2 T08148	proline-rich myros
3	116.5	12.1	676	1 EDBE23	immediate-early pr
4	116	12.0	862	2 T46289	hypothetical prote
5	113.5	11.8	422	2 S32357	glial growth facto
6	111	11.5	801	2 T29018	hypothetical prote
7	107	11.1	2142	2 B35098	MHC class III hist
8	106.5	11.0	356	1 W0RU2H	homeotic protein H
9	106	11.0	325	2 A55558	albumin D-box bind
10	104.5	10.8	3938	2 T42761	Baasoon protein -
11	103.5	10.7	924	2 S72923	gene Lf3 protein -
12	103	10.7	505	2 S72273	actin-depolymerizi
13	103	10.7	1611	2 T38236	hypothetical prote
14	103	10.7	1952	2 T48814	hypothetical prote
15	103	10.7	3436	2 S55659	tegument protein 6
16	102.5	10.6	676	1 EDBE22	immediate-early pr
17	102.5	10.6	862	2 S51493	major nitrogen reg
18	102	10.6	317	2 T19143	hypothetical prote
19	102	10.6	365	2 A34894	hepatic transcript
20	101	10.5	1870	2 S37671	MHC class IIT hist
21	101	10.5	1872	2 S36152	MHC class III hist
22	101	10.5	2240	2 T30517	probable multi-dom
23	100	10.4	574	2 F75356	serine/threonine p
24	100	10.4	592	2 D70863	hypothetical prote
25	99	10.3	395	2 H75457	hypothetical prote
26	99	10.3	1334	2 T50568	probable multi-dom
27	99	10.3	1487	1 EDBEF6	155k transcripion
28	98	10.2	1487	1 EDBBE1	immediate-early pr
29	97.5	10.1	384	2 S51796	vasodilator-stimul

30	97.5	10.1	405	2 G70829	probable molybdenu
31	97.5	10.1	896	2 S36326	clathrin assembly
32	97.5	10.1	915	2 S36327	clathrin assembly
33	97	10.1	353	2 S36249	112b protein - Pe
34	96.5	10.0	555	2 T30349	structural protein
35	96.5	10.0	708	2 D96711	hypothetical prote
36	96.5	10.0	1255	2 T31065	diaphanous protein
37	96.5	10.0	1942	2 T42730	Baasoon protein -
38	96	9.9	548	2 S52735	CM1R protein - mo
39	96	9.9	581	2 T22341	hypothetical prote
40	96	9.9	3164	1 MMBE6	UL36 protein - hum
41	95.5	9.9	316	2 T19291	hypothetical prote
42	95.5	9.9	903	2 T00705	N-chimerin homolog
43	95	9.8	302	2 T32711	hypothetical prote
44	95	9.8	607	2 A43776	drebrin E2 - chick
45	95	9.8	2187	2 T30826	nascent polypeptid

## ALIGNMENTS

RESULT 1  
SKXLAG  
dermal gland protein APEG precursor - African clawed frog (fragment)  
C:Species: Xenopus laevis (African clawed frog)  
C>Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 16-Jun-2000  
C:Accession: S07498; A34140  
R:Gmahl, M.; Berger, H.; Thallhammer, J.; Krell, G.  
FEBS Lett. 260, 145-148, 1990  
A>Title: Dermal glands of Xenopus laevis contain a polypeptide with a highly repetitive  
A:Reference number: A34140; MUID:90127399; PMID:2298293  
A:Accession: S07498  
A:Molecule type: mRNA  
A:Residues: 1-416 <GMA>  
A:Cross-References: EMBL:X51394; NID:664547; PIDN:CAA35759.1; PID:G1334643  
A>Note: the authors translated the codon TTC for residue 8 as Pro  
C:Superfamily: dermal gland protein APEG; trefol1 homology  
C:Keywords: skin; tandem repeat  
F:1-20/Domain: signal sequence (fragment) #status predicted <SIG>  
F:21-416/Product: dermal gland protein APEG #status predicted <MAT>  
F:23-344/Region: alanine/glutamic acid/glycine/proline-rich repeats  
F:349-389/Domain: trefol1 homology <TRF>  
F:349-375,359-374,369-386/Dileulide bonds: #status predicted

Query Match  
Beet Local Similarity 27.5%; Score 124; DB 1; Length 416;  
Matches 46; Conservative 10; Mismatches 57; Indels 54; Gaps 7;

QY 12 DAPAPTPC-----VPAECFDLVRHCACGLRTPPKPAGASSPAPRTALPQESVG 64  
DB 153 EAPAPAPAEVEAPAPAPAE-----EAPAPAPAEAPAPAPAPAEAPAPAPAE 198  
QY 65 AGAGGALPLPGLFGAPALGLALVITALVGLVSWRRORRLGASSAEP-DGDKDA 123  
DB 199 PAEGAPAPAP-----APA-----EGAPAPAPAEAPAPAE 228  
QY 124 PEPLD-KYIIISPGISDAPAPPPGDPPTTPPGHSVPVATLIG 169  
DB 229 PAPAGAPAPAPAPAEAPAPAPAE-PAEGAPAPAPAEAPAPAPAE 274

RESULT 2  
T08148  
proline-rich myrosinase-binding protein homolog - rape (fragment)  
N:Alternate names: myrosinase-binding protein related protein  
C:Species: Brassica napus (rape)  
C>Date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 08-Oct-1999  
C:Accession: T08148  
R:Raipalensu, J.; Falk, A.; Ek, B.; Rask, L.  
Eur. J. Biochem. 243, 605-611, 1997  
A>Title: Myrosinase-binding proteins are derived from a large wound-inducible and repeti  
A:Reference number: Z16379; MUID:97210758; PMID:9057822  
A:Accession: T08148





```

Db      170 RPRRVETIALLDLTETQVKWFQNRMRKHQKOTQHRREPPDGBACPGALIEDICDPAEEPA 229
QY      133 LSPGISDATAPAV-----PP--PG-----EDPGTTPPG 158
         ||| : || || || | : ||
Db      230 ASFGCPASAPAAWEACHPREVPPGSLSDPPPLAVRLGAGASSPG 276

```

RESULT 9  
A55558  
albumin D-box binding protein - human  
C/Species: Homo sapiens (man)  
C/Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 01-Dec-2000  
C/Accession: A55558; G02887  
R/Katib, Z.A.; Inaba, T.; Valentine, M.; Look, A.T.  
Genomic 23,344-351,1994  
A/Title: Chromosomal localization and cDNA cloning of the human DBP and TEF genes  
A/Reference number: A55558; MUID:95137580; PMID:7835883  
A/Accession: A55558  
A/Status: preliminary  
A/Molecule type: mRNA  
Residues: 1-325 <KHA>  
Cross-references: GS:U06936; NID:g606798; PID:AAA81374.1; PID:g606799

submitted to the EMBL Data Library, February 1996

A:Accession: G02887

A:Reference number: H01797

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-178, 'S', 180-325 <MUE>

A:Cross-references: EMBL:U48213, NID:g1208753, PID:g1208754

C:GeneticCB:

A:Gene: GDB:DBP

A:Cross-references: GDB:128840, OMIM:124097

A:Map position: 4q12-4q12

A:Introns: 47/1; 184/1; 254/3

[illegible]

```

RESULT 10
T42761      Bassoon protein - rat
N/Alternate names: brain-specific synapse-associated protein
C/Species: Rattus norvegicus (Norway rat)
C/Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 21-Jul-2000
C/Accession: T42761
R/Dieck, S.; Sammarti-Vila, L.; Langnaese, K.; Richter, K.; Kindler, S.; Soyke, A.; Wex,
J. Cell Biol. 142, 499-509, 1998
A/Title: Bassoon, a novel zinc-finger CAG/Glutamine-repeat protein selectively localized
A/Reference number: Z22249; MUID:98345363; PMID:9679147
A/Accession: T42761
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-3938 <DIE>
A/Cross-references: EMBL:Y16563; NID:G3413503; PIDN:CAA76287.1; PID:G3413504

```

A:Experimental source: strain Sprague Dawley; brain  
C:Function:  
A:Description: may be involved in cytomatrix organization at the site of neurotransmitter  
A:Note: component of the presynaptic cytoskeleton  
C:Keywords: coiled coil; zinc finger

Query Match	10.8%;	Score 104.5;	DB 2;	Length 3938;
Best Local Similarity	26.8%;	Pred. No. 11;		
Matches	53;	Conservative	13;	Mismatches 73;
				Indels 59;
				Gaps 10

```

OY  GPP---SLGRDADPATPC-VPAEQFDLVHCVACGLRPPRPKPGAGSPARTALOP 59
Db  272 GPPQAEAAARTISVPGCTQATAPPE-----VGRVSPQPLSTKSTI-EP 314
OY  60 QESVGAGAGAAALPLP-----GLLFGAPALLGLIALVLALVIVGLVSWRRQ 105
Db  315 RPPAGEAGKGSATTVTSVGLGAABQQTGGITGLRPLGASL---LTQASTLSVQP----- 366
OY  106 RLRLGASSAAEPDGDKDAPEPLDKYIIISPGISDATAPAMPPP---GBDQTTTPGHVSP 162
Db  367 ----EADTQOQPEPSKGP-----KIV-----FSDASKENAGPRPGSGGPGPTPGAKTEP 413
OY  163 VPATLGSSTELVTTKAG 180
Db  414 GPRTPGSGGALAKTGG 431

```

RESULT 11  
S27923  
Gene lF3 protein - human herpesvirus 4  
C.Species: human herpesvirus 4, Epstein-Barr virus  
C.Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #ext\_change 31-Jan-2000  
C.Accession: S27923  
R.Parker, B.D.; Bankier, A.; Satchwell, S.; Barrell, B.; Farrell, P.J.  
submitted to the EMBL Data Library, August 1990  
A.Description: Sequence and transcription of Raji Epstein-Barr virus DNA spanning the B  
A.Reference number: S27923  
A.Accession: S27923  
A.Status: Preliminary  
A.Molecule type: DNA  
A.Residues: 1-924 <PAR>  
A.Cross-references: EMBL:M35547, NID:G330420, PID:AAA5896.1, PID:G330421  
C.Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology/;

[illegible]

RESULT 12  
S72273  
actin-depolymerizing protein N-WASP, brain - bovine  
C|Species: Bos primigenius taurus (cattle)  
C|Date: 23-Apr-1998 #sequence\_revision 01-May-1998 #text\_05-Nov-1999  
C|Accession: S72273  
R|Miki, H.; Miura, K.; Takenawa, T.  
EMBO J. 15, 5326-5335, 1996

A:Title: N-WASP, a novel actin-depolymerizing protein, regulates the cortical cytoskeleton  
 A:Reference number: S72273; MUID:97050838; PMID:8895577  
 A:Accession: S72273  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-505 <MIK>  
 A:Cross-references: EMBL:D67066; NID:g1644231; PDB:BA11082.1; PID:d1011742; PID:g16442  
 A:Experimental source: brain

Query Match 10.7%; Score 103; DB 2; Length 505;  
 Best Local Similarity 23.6%; Pred. No. 2;  
 Matches 41; Conservative 4; Mismatches 45; Indels 84; Gaps 7;

QY 5 PSLGRDAPAPTPC-VPAECFDLVRHCACGLLRTPPKPA-----GASSAPARTALOP 59  
 DB 301 PPPAGRGAPPPPSAPPTNA-----PPPPPSREVGAPPPPPNMP 345  
 QY 60 QESVAGAGEAALPLPGLLFGAPALLGLALVTLVGLVSWRRORRLGASSAEPDQ 119  
 DB 346 -----PLPAL-----PSSAPSG 357  
 120 DKDAPEPLDKVIIISPGISDATAPAMPPEGDEGTTTP-----GHSVPVPA 165  
 DB 358 PPPPPPL-----SVSGSVAPPPPPPPPPPPGLPSDGDHQPPTPA 403

RESULT 13

hypothetical protein SPAC23A1.17 - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C:Accession: T38236  
 R:Murphy, L.; Harris, D.; Wood, V.; Barrell, B.G.; Rajandream, M.A.  
 submitted to the EMBL Data Library, February 1998  
 A:Reference number: Z21780  
 A:Accession: T38236  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1611 <MUR>  
 A:Cross-references: EMBL:AL021813; PDB:1CA16991.1; GSPDB:GN00066; SPDB:SPAC23A1.17  
 A:Experimental source: strain 972h-; cosmid c23A1  
 C:Genetics:  
 A:Gene: SPDB:SPAC23A1.17  
 A:Map position: 1

Query Match 10.7%; Score 103; DB 2; Length 1611;  
 Best Local Similarity 24.6%; Pred. No. 6;  
 Matches 44; Conservative 16; Mismatches 71; Indels 48; Gaps 7;

DB 13 AAPTPCVPAEFGEDLVRHCACGLLRTPPKPAAGASSAPARTALOPQESVAGAGEAAL 72  
 DB 1078 AMSGAPVPA-----PSGITPPVKPSVAAPVPPKPSVAVPP--VAPSGAPV 1123  
 QY 73 PLPGLI-----FGAPALLGLALVTLVGLVSWRRORRLGASSAEPDGDQDAE 125  
 DB 1124 PKPSVAPAPVPPVSGAPV-----PKPSVAPAPVAPSG--AP- 1159  
 QY 126 PLDKVIIISPGI---SDAPAPAMPPEGDEGTTTPPGHSVVPATLGLSTELVTTKAGP 181  
 DB 1160 PVPKDSVAPAPVAPSSGIPVPPKPAAGVPPVPPSEADPPKPSVGVGVPPPPSTAPP 1218

RESULT 14

hypothetical protein 15E6.220 [imported] - Neurospora crassa  
 C:Species: Neurospora crassa  
 C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 05-May-2000  
 C:Accession: T48814  
 R:Schulte, U.; Aism, V.; Hohenseel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,  
 submitted to the Protein Sequence Database, April 2000  
 A:Reference number: Z24541  
 A:Accession: T48814  
 A:Status: preliminary

A:Molecule type: DNA  
 A:Residues: 1-1952 <SCH>  
 A:Cross-references: EMBL:AL353822; GSPDB:GN00112; NCSP:15E6.220  
 A:Experimental source: cosmid contig 15E6; strain 74  
 C:Genetics: NCSP:15E6.220  
 A:Gene: NCSP:15E6.220  
 A:Map position: 2  
 A:Introns: 281/3

Query Match 10.7%; Score 103; DB 2; Length 1952;  
 Best Local Similarity 29.0%; Pred. No. 7.2;  
 Matches 38; Conservative 15; Mismatches 62; Indels 16; Gaps 4;

QY 41 PPPKPAAGSSAPARTALOPQESVAGAGEAALPLPGLLFGAPALLGLALVTLVGLV 99  
 DB 1812 PAPSAAQSVAPAPVSTVPAATVAPASTVAAAPPTVTATAPALSAATNPAPV----- 1867  
 QY 100 SWRRORRLRGASSAEPDGDQDAEPLDKVIIIS-----PCISDATAPAMPPEGDEG 154  
 DB 1868 --SQPHOITGQAPAQOQRPAPAOAPATPTTSAAPPPTLAPPPPPPPTEDPPP 1925  
 QY 155 TPPGHSVPVPA 165  
 DB 1926 PPP-----PPPA 1932

RESULT 15

legumem protein 64 - equine herpesvirus 2  
 C:Species: equine herpesvirus 2  
 C:Date: 27-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 08-Oct-1999  
 C:Accession: S55659  
 R:Telford, E.A.R.; Watson, M.S.; Aird, H.C.; Perry, J.; Davison, A.J.  
 J. Mol. Biol. 249, 520-528, 1995  
 A:Title: The DNA sequence of equine herpesvirus 2.  
 A:Reference number: S55594; MUID:95302501; PMID:7783207  
 A:Accession: S55659  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-3436 <TEL>  
 A:Cross-references: GB:U020824; NID:g695172; PDB:1AC13852.1; PID:g695237  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1995

Query Match 10.7%; Score 103; DB 2; Length 3436;  
 Best Local Similarity 25.0%; Pred. No. 12;  
 Matches 55; Conservative 12; Mismatches 51; Indels 102; Gaps 11;

QY 2 RRGPSRLGRD-----APAPT-----PCVPAECFDL-----VRHCYAC 35  
 DB 3178 RLGERGRERREVEVAADAPRDYSRVPVPRREPSCGPRSPARDDTLGGAGSREKVS 3237  
 QY 36 G-----LLRPPKPKAGASSAPARTALOPQESVAGAGEAALPLPGLLFGAPALLGLA 88  
 DB 3238 GGRPRVOLSRSPKPRPASQVQGR-----EEVGFSPGRAR----- 3274  
 QY 89 LVLALVGLVSWRRORRLRGASSAEPDGDQDAEPLDKVIIIS----- 135  
 DB 3275 -----KGGSTAHAP-----PE-TDTRADYIEPKSGVAGSGSPE 3307  
 QY 136 -----GISDATAPAMPPEGDEGTTTPPGHSVVPATLGLST 171  
 DB 3308 KKQGAABAEAPVAP-PRGEGAAQAEAPGEGTPEPAS-IGKT 3345

Search completed: June 23, 2003, 15:16:43  
 Job time : 49 secs

**THIS PAGE BLANK (USPTO)**